Available online at www.sciencedirect.com

SCIENCE DIRECT

Discrete Applied Mathematics 154 (2006) 1028–1031

DISCRETE APPLIED MATHEMATICS

www.elsevier.com/locate/dam

Note

Latin squares with bounded size of row prefix intersections

Grzegorz Malewicz[1]

Department of Computer Science, University of Alabama, USA

Received 7 September 2004; accepted 8 November 2005

Abstract

A latin square is a matrix of size $n \times n$ with entries from the set $\{1, \ldots, n\}$, such that each row and each column is a permutation on $\{1, \ldots, n\}$. We show how to construct a latin square such that for any two distinct rows, the prefixes of length h of the two rows share at most about h^2/n elements. This upper bound is close to optimal when contrasted with a lower bound derived from the Second Johnson bound [6].

Keywords: Combinatorial design theory; Coding theory; Graph theory; Latin squares; Graph coloring; Second Johnson bound; Youden squares

1. Construction

A *latin square* [3] is a matrix of size $n \times n$ with entries from the set $\{1, \ldots, n\}$, such that each row and each column is a permutation on $\{1, \ldots, n\}$. The number n is called the *order* of the latin square. We construct a latin square S of order $n = q^2 + q + 1$, for any q prime power. This latin square has additional properties. For any two distinct rows, their prefixes of length h share at most $q + \lceil h/q \rceil^2$ elements. The upper bound is contrasted with a lower bound obtained from the Second Johnson bound stating that worst-case intersection must be at least $h^2/n - 4$. Our construction extends the recursive construction of a system of schedules $\mathcal{G}_n$ from [9]. Specifically, we show that we can rearrange the elements of each schedule so that a matrix that contains the rearranged schedules as rows is a latin square.

We report a few standard facts from Design Theory [1,5]. Pick any prime power q. Let $GF(q)^3$ be the vector space of dimension 3 over the finite field of order q. It is known that there are exactly $n = q^2 + q + 1$ distinct subspaces of dimension 2 in this space. We call them planes and denote $P_1, \ldots, P_n$. There are exactly n distinct subspaces of dimension 1. We call them lines and denote $L_1, \ldots, L_n$. For any plane, there are exactly $q + 1$ distinct lines included in the plane. For any line, there are exactly $q + 1$ distinct planes that include the line.

Our construction begins by coloring a bipartite graph. Let $\mathcal{G}$ be a bipartite graph on $2n$ nodes. The n nodes on the left correspond to the n lines, and the n nodes on the right correspond to the n planes. There is an edge between left node i and right node j if and only if line L_i is included in plane P_j. This graph can be constructed in time $O(n \log n + nq) = O(n^{3/2})$ when q is prime, because for any left node i, in order to find all adjacent right nodes, it is sufficient to find at most 2 multiplicative inverses in $GF(q)$ and perform $O(q)$ other arithmetic operations in $GF(q)$ (details can be found for example in [8] Fig. 2). Specifically, for any left node i, we can determine all its right neighbors

[1] Present address: Google Inc., Department of Engineering, 1600 Amphitheater Parkway, Mountain View, CA 94043, USA.
E-mail address: malewicz@google.com.

0166-218X/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2005.11.002

$R(i)$, and for any right node j, we can determine all its left neighbors $L(j)$. From the earlier discussion the graph is $q+1$ regular; each node on the left is connected to $q+1$ nodes on the right, and each node on the right is connected to $q+1$ nodes on the left. Malewicz et al. [9] considered this graph and observed that it has a perfect matching. We extend this observation, and explore ramifications of the extension. A classical result by Konig [7] yields that edges of this graph can be colored with $q+1$ colors such that no node has two adjacent edges colored with the same color. For any left node i, let $p_{i,c}$ be the right node to which an edge of color c leads from i, $1 \leqslant i \leqslant n$, $1 \leqslant c \leqslant q+1$. This coloring can be found in time $O(n^{3/2} \log n)$ using the algorithm of Cole and Hopcroft [2] (see [12] for recent results on bipartite graph coloring). By the construction, for any c, the numbers $p_{1,c}, p_{2,c}, \ldots, p_{n,c}$ are district numbers from the set $\{1, \ldots, n\}$, because any node on the right has a single adjacent edge of any given color c. For any left node i, the numbers $p_{i,1}, p_{i,2}, \ldots, p_{i,q+1}$ are distinct numbers from the set $\{1, \ldots, n\}$, because they correspond to the $q+1$ distinct nodes on the right connected to the left node i. The resulting matrix $(p_{i,c})$ is referred to as Youden square [4,11,13].

We pick a c, and show how to sequence the lines of each of the n planes $P_{p_{1,c}}, \ldots, P_{p_{n,c}}$ so that any ith line in a sequence is distinct across the n sequences. We observed that the planes $P_{p_{1,c}}, \ldots, P_{p_{n,c}}$ are distinct, so they constitute all the n planes of $GF(q)^3$. By the construction of the graph $\mathcal{G}$, the plane $P_{p_{i,c}}$ contains line L_i and q other lines. We consider a bipartite graph $\mathcal{G}_c$ with n nodes on the left that correspond to the planes, and n on the right that correspond to the lines. We link a plane $P_{p_{i,c}}$ on the left with all lines on the right that the plane includes, except for line L_i i.e., to the set $L(p_{i,c}) \backslash \{i\}$ of nodes on the right. This graph $\mathcal{G}_c$ can be constructed in time $O(n^{3/2})$ because sets $L(i)$ are already known. This graph has the property that each node on the left is linked to exactly q nodes on the right. In addition, each line on the right is included in exactly $q+1$ planes, and exactly one of them is not linked to the line, so any node on the right is linked to exactly q nodes on the left. Hence the graph $\mathcal{G}_c$ is q-regular, and thus its links can be colored in time $O(n^{3/2} \log n)$ using q colors. Let $p^c_{i,d}$ be the right node to which an edge of color d leads from i, $1 \leqslant i \leqslant n$, $1 \leqslant d \leqslant q$. So $p^c_{i,1}, p^c_{i,2}, \ldots, p^c_{i,q}$ are the indices of all the lines other than L_i that are in the plane $P_{p_{i,c}}$. Moreover, by the property of the coloring, for any color d, the numbers $p^c_{1,d}, p^c_{2,d}, \ldots, p^c_{n,d}$ are all distinct numbers from the set $\{1, \ldots, n\}$.

We are now ready to construct an n by n matrix S that is a latin square. The first column of the matrix lists numbers from 1 to n in that sequence starting from the top row. The subsequent $q(q+1)$ columns are partitioned into $q+1$ groups of q consecutive columns each; group number c, $1 \leqslant c \leqslant q+1$, contains q columns. Column number d, $1 \leqslant d \leqslant q$, of this group lists numbers $p^c_{1,d}, p^c_{2,d}, \ldots, p^c_{n,d}$ in that sequence starting from the top row. As a result, each column of the matrix is a permutation. To see why each row is also a permutation we restate the argument presented in [9]. Row i contains i followed by $q+1$ groups of q elements each. Each group c contains indices of all lines contained in the plane $P_{p_{i,c}}$ excluding the index of line L_i that is also contained in the plane. But any two distinct planes in $GF(q)^3$ intersect at a single line, so each group contains exactly q elements that are different from the elements that any other group of this row contains. So the total number of distinct elements in this row is $1+(q+1)q = n$, and these elements are from $\{1, \ldots, n\}$. Thus the row must be a permutation. We summarize this construction in the following theorem. An example of matrix S for $q = 3$ is given in Fig. 1.

1	2 3 13	4 9 12	7 6 11	8 5 10
2	10 8 5	13 6 4	1 11 9	3 7 12
3	5 7 12	6 8 1	9 4 2	10 13 11
4	12 9 2	11 3 8	6 5 13	7 10 1
5	9 11 1	2 7 6	10 13 12	4 3 8
6	13 5 4	12 11 10	2 1 3	9 8 7
7	6 2 11	9 10 3	8 12 1	13 4 5
8	7 13 9	5 1 11	3 10 6	2 12 4
9	3 10 6	7 12 5	11 8 4	1 2 13
10	4 1 7	8 13 9	12 3 5	11 6 2
11	8 4 3	10 5 2	13 9 7	12 1 6
12	1 6 8	3 2 13	4 7 10	5 11 9
13	11 12 10	1 4 7	5 2 8	6 9 3

Fig. 1. A latin square of order 13 with controlled size of intersection of prefixes of any two distinct rows.

Theorem 1.1. *For any prime power q, let $n = q^2 + q + 1$. Then the matrix S of size n by n is a latin square. When q is a prime, S can be constructed in time $O(n^2 \log n)$.*

A bound on the size of intersection of prefixes of any two distinct rows of S follows from Theorem 5 of [9].

Theorem 1.2 (*Makewicz et al. [9]*). *For any $0 \leqslant i \leqslant q + 1$, the cardinality of the intersection of the sets of the leftmost $1 + iq$ elements of any two distinct rows of S is 0 when $i = 0$, 1 when $i = 1$, and at most $q + i^2$ when $2 \leqslant i \leqslant q + 1$.*

Briefly speaking, this theorem results from the fact that any two distinct planes intersect at a single line (hence the summand i^2), and that any two rows were formed from certain planes and only one pair of these planes may be the same, because the planes for the former row are all planes that include a line and for the later row that include a different line (hence the summand q); each first element is distinct (hence the bound is 0 when $i = 0$), and the plane that formed the first group of each row is distinct across rows (hence the bound is 1 when $i = 1$).

The bound on pairwise intersection of rows of S is close to optimal when contrasted with the Second Johnson bound [6]. A u-(t, h, λ) *packing design* [10,14] is a family of subsets $S_1, \ldots, S_n$ of the set $\{1, \ldots, t\}$ with the property that each $|S_i| = h$ and any set of u elements of $\{1, \ldots, t\}$ is a subset of at most λ of the S_i. The subsets S_i are called *blocks*. The *packing number* $D_\lambda(t, h, u)$ is the maximum number n of blocks in any u-(t, h, λ) packing. Johnson [6] gave an upper bound on the packing number.

Theorem 1.3 (*Second Johnson bound, [6]*). *The packing number $D_1(t, h, u)$ is bounded by*

$$D_1(t, h, u) \leqslant \frac{t(h + 1 - u)}{h^2 - (u - 1)t}$$

This bound lets us immediately state that there must be a pair of distinct rows whose prefixes have large intersection.

Corollary 1.4. *For any matrix of size n by n with elements from $\{1, \ldots, n\}$, there are two rows such that their prefixes of length h have at least*

$$h^2/(n - 1) - (h + n - 1)/(n - 1)$$

elements in common.

Proof. Take any such matrix and suppose that the intersection of prefixes of length h of any two distinct rows is at most $u - 1$. The n prefixes form n blocks of a u-$(n, h, 1)$ packing design. From the Second Johnson bound we get that $n \leqslant n(h + 1 - u)/(h^2 - (u - 1)n)$, and so $u \geqslant h^2/(n - 1) - h/(n - 1)$. Hence the intersection $u - 1$ must be at least $h^2/(n - 1) - h/(n - 1) - 1$, for some pair of prefixes of distinct rows. □

The matrix S is in semi-standard form. The first column lists numbers from 1 to n in that sequence starting from the top row, while the first row may not. We may reshuffle columns 2–n to bring it to the standard form (to get the identity permutation in the first row), but then we may loose control over the size of intersection of prefixes of rows.

Acknowledgements

The author thanks Charles Colbourn, Donald Preece and Neil J.A. Sloane for discussions on combinatorial design theory, and Christos Kaklamanis for a discussion on graph coloring. The result presented here was developed when the author was a Ph.D. student. The author thanks Alex Russell and Alex Shvartsman for guidance in research during that time.

References

[1] C.J. Colbourn, J.H. Dinitz, CRC Handbook of Combinatorial Designs, CRC Press, Boca Raton, FL, 1996.
[2] R. Cole, J. Hopcroft, On edge coloring bipartite graphs, SIAM J. Comput. 11 (3) (1982) 540–546.

[3] J. Denes, A.D. Keedwell, Latin Squares and their Applications, Academic Press, New York, 1974.
[4] H.O. Hartley, S.S. Shrikhande, W.B. Taylor, A Note on Incomplete Block Designs with Row Balance, Ann. Math. Statist. 24 (1) (1953) 123–126.
[5] D.R. Hughes, F.C. Piper, Design Theory, Cambridge University Press, Cambridge, 1985.
[6] S.M. Johnson, A New Upper Bound for Error-Correcting Codes, IEEE Trans. Inform. Theory 8 (1962) 203–207.
[7] D. Konig, Graphok 'es alkalmaz'asuk a determin'asok 'es a halmazok elm'elet'ere, Mathematikai 'es Term'eszettudom'anyi ' Ertesito 34 (1916) 104–119.
[8] G. Malewicz, A. Russell, A.A. Shvartsman, Distributed Cooperation During the Absence of Communication, 14th International Conference on Distributed Computing, 2000, pp. 119–133.
[9] G. Malewicz, A. Russell, A.A. Shvartsman, Optimal Scheduling for Disconnected Cooperation. Eighth International Colloquium on Structural Information and Communication Complexity, 2001, pp. 259–274.
[10] W.H. Mills, R.C. Mullin, Coverings and packings, in: J.H. Dinitz, D.R. Stinson (Eds.), Contemporary design theory: a collection of surveys, Wiley, New York, 1992, pp. 371–399.
[11] D.A. Preece, Fifty Years of Youden Squares: a Review, Bull. Inst. Math. Appl. 26 (1990) 65–75.
[12] A. Schrijver, Bipartite Edge Coloring in $O(\Delta m)$ Time, SIAM J. Comput. 28 (3) (1998) 841–846.
[13] C.A.B. Smith, H.O. Hartley, The Construction of Youden Squares, J. Roy. Statist. Soc. Ser. B (Methodol.) 10 (2) (1948) 262–263.
[14] D.R. Stinson, Packings, in: C.J. Colbourn, J.H. Dinitz (Eds.), Packings, in CRC Handbook of Combinatorial Designs, CRC Press, Boca Raton, FL, 1996, pp. 409–413.